INTELLIGIBILITY OF SELECTED CONSONANT SOUNDS DISTORTED BY INFINITE PEAK CLIPPING

By

MAURICE JOSEPH

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

December, 1966

To My Wife

ACKNOWLEDGMENTS

The author expresses sincere appreciation to his advisor and chairman, Dr. Donald Dew, for guidance throughout the planning and execution of this study, and to members of the dissertation committee for suggestions and encouragement.

Special appreciation is expressed to Dr. Harry Hollien, and the faculty, staff, and students of the Communication Sciences Laboratory of the University of Florida for cooperation in all phases of the study.

In addition, the author gratefully acknowledges Dr. A. E. Brandt's advice and aid in statistical treatment of the data.

The author wishes especially to thank his wife, Pauline, for her constant encouragement, inspiration, and untiring aid at all times.

Finally, the author expresses gratitude to the Vocational Rehabilitation Administration for financial aid through a Special Rehabilitation Research Fellowship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

BACKGROUND AND STATEMENT OF PROBLEM

Introduction

Experimental studies have demonstrated that word lists retain intelligibility when subjected to distortion by infinite peak clipping (Licklider and Pollack, 1948; Licklider, Bindra, and Pollack, 1948). These findings have been interpreted to mean that the temporal pattern of zero-axis crossings of the speech waveform provides information that is in itself sufficient for perception of speech. A corollary inference is that clues provided by relative amplitude variations in the speech waveform may be discarded without degradation of intelligibility. These interpretations of the effects of infinite peak clipping are of considerable theoretical importance: to the extent that they are true they describe an invariant factor in speech intelligibility. However, there are two reasons to suspect that these generalizations may need qualification. One reason is that the speech sample used in the early studies may have provided contextual clues to word perception in addition to the frequency and amplitude clues inherent in the speech wave. Thus, with multiple clues

operating, the role of one physical parameter is more difficult to assess. A second reason to question the generalizations about zero-axis crossing and relative amplitude clues is that these clues may assume different importance for different types of speech. Factors that affect intelligibility of word lists may differ from those that affect individual sounds; moreover, it is possible that these factors may assume different importance among different individual speech sounds.

It may be argued that speech communication generally involves units larger than individual speech sounds, and that word intelligibility provides a closer approximation to intelligibility of conversational speech. It may be true that whole words, with contextual clues free to operate, have higher face validity as representative of speech in general. However, speech communication situations do exist--in military communications, for instance--in which identification of specific speech sounds may be essential. Moreover, fundamental generalizations about factors that determine intelligibility of speech should be capable of verification on molecular levels (figuratively) such as individual phonemes, consonant clusters, and syllables, as well as on the molar level of words and continuous discourse.

The focus of this study is on the individual speech sound. Specifically, the present study represents an attempt to determine the effects of infinite peak clipping on

intelligibility of specific consonants. It also represents an attempt to minimize the opportunity for contextual clues to operate.

Infinite Peak Clipping

Definition

Peak clipping is a form of distortion in which amplitude is prevented from exceeding a certain maximum limit. Amplitude is held constant during times when, in a linear system, it would exceed the limiting level. When the positive and negative-going parts of the speech wave are equally limited, the effect is said to be "symmetrical peak clipping." If a wave thus limited is reamplified so that the clipped output wave equals the amplitude of the peaks of the original wave (before clipping), and if this operation is continued through several successive stages of clipping and reamplification until the waveform is approximately rectangular in shape, the operation is called "infinite peak clipping." Line B of Figure 1 illustrates the result of infinite peak clipping on the waveform shown on Line A. Two characteristics of the clipped wave may be noted on this figure: the points (in time) that correspond to axis crossings in the undistorted wave are the same in the clipped wave, but the amplitudes have only two values, a single maximum excursion above and below the axis. In addition, all indications of differences in rise-time or fall-time, or of changes that did not cross

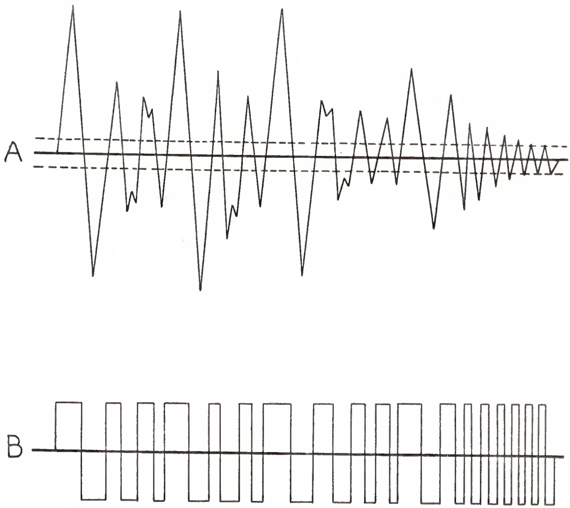


Figure 1. A comparison of a complex wave before and after infinite peak clipping (after Licklider, Bindra, and Pollack, 1948).

A. Undistorted
B. Same wave following infinite peak clipping

the axis, are eliminated in the clipped wave.

Relevant literature

Effects on intelligibility. The most important study on the intelligibility of speech distorted by infinite clipping was that of Licklider and Pollack (1948). Phonetically balanced word lists recorded on discs by one talker were subjected to infinite peak clipping and presented to five listeners in a series of twenty-five sessions. Scores averaged approximately sixty-eight percent in the first listening session, but improved in succeeding sessions. Intelligibility scores of over ninety percent were achieved by the twenty-fifth session. The authors felt that this increase was independent of the listeners' increased familiarity with PB word lists but instead represented learning to understand clipped speech. This suggests that statements regarding the intelligibility of clipped speech should specify the observers' familiarity with clipped speech.

The effects of peak clipping in combination with other forms of distortion were also studied. Integration of the waveform, the equivalent of a low-pass filter slope of six dB per octave, severely degrades intelligibility of PB-50 word lists, if the integration occurs prior to clipping. Differentiation of the waveform, the equivalent of a high-pass filter slope of six dB per octave, performed before clipping, results in intelligibility slightly better than

with clipping alone. Integration or differentiation of the waveform following the clipping had no effect on intelligibility (Licklider and Pollack, 1948). When combined with distortion by reduction of sound duration, subsequent clipping had little effect on the intelligibility of vowels, but did reduce the intelligibility of consonants (Ahmend and Fatehchand, 1959) in CVC and VC syllables.

Spectral effects. In attempting to predict or to explain the effects of infinite peak clipping under various conditions, it would be helpful to know the effects of clipping on the spectrum of speech. Unfortunately, the information available is scarce and somewhat equivocal. Only one report gives empirical measurements: Licklider, Bindra, and Pollack (1948), reporting the results of an intensity-frequency-time spectral analysis of the word "shoe-bench," both undistorted and peak clipped, state ". . . although many of the details of the pattern are changed by infinite peak clipping, the general plan of the terrain is by no means rendered unrecognizable. The main concentrations of low-frequency and of high-frequency energy are still in the same places despite the rearrangement of minor peaks." This report, even while minimizing the effects of clipping on the spectrum, does support the contention that changes do occur in the spectrum under infinite peak clipping.

Two theoretical papers, analyzing the clipping process

mathematically, exhibit similar ambiguity. Velichkin (1962) concludes ". . . It is evident that amplitude clipping causes a broadening of the speech spectrum, but this broadening is slight, even with maximum clipping." Mathematical calculations by Dukes (1954) confirmed the likelihood that clipping would not seriously reduce the intelligibility of speech, but he adds, ". . . it is clear that the results have significance in respect of very long samples . . . With this formulation nothing can be deduced about the intelligibility of individual sounds, except of course that large deviations below the average must be relatively infrequent." He points out that in applying his formulas some of the values obtained represent averages over all possible unvoiced and voiced sounds.

The most serious shortcoming of the spectral information available, whether empirical or theoretical, is that it is based on relatively long-term averaged values. Apparently spectral changes, however small, do occur under infinite peak clipping, but are obscured by averaging over time. Specific moment-to-moment changes are not yet known in detail; this makes it difficult to predict the effects on brief speech sounds such as consonants.

Implications of previous studies on peak clipping

Two related implications of the Licklider and Pollack (1948) study have been considered of major importance. These concern the relative importance of dynamic amplitude variation

and of the temporal pattern of zero-axis crossings in the speech waveform.

Amplitude clues. Since the speech wave subjected to infinite peak clipping assumes a dichotomous value of amplitude, the almost infinite range of possible amplitude values present in the undistorted wave is reduced to one value of maximum and minimum excursion. This, in effect, means that information regarding relative amplitudes in the speech wave is discarded. The implication drawn is that ". . . the variations in intensity from moment to moment appear not to be basic cues for the recognition of words," and that "the so-called dynamic characteristics of speech are not of vital importance for intelligibility. It is apparently just as well to reproduce all the fundamental speech sounds (or what is left of them after clipping) at the same intensity level as it is to preserve their normal intensities" (Licklider and Pollack, 1948).

Zero-axis crossings. It has been reasoned that since infinite peak clipping eliminates information regarding relative amplitudes, as well as on- and off-slope and wave shape, all that remains in the clipped wave is the temporal pattern of zero-axis crossings. The inference drawn is that the axis crossing information is sufficient to provide speech intelligibility (Licklider, 1950). It should be noted that this is negative evidence based upon the elimination of other possibilities.

Black and Hixson (1959) were unable to demonstrate a direct relationship between density of zero-axis crossings and intelligibility.

The implications of the Licklider and Pollack study may be paraphrased as follows:

1) The information provided by the temporal pattern of zero-axis crossings is necessary and sufficient for intelligibility of speech, and

2) The information provided by the dynamic pattern of amplitude variations is not necessary for intelligibility of speech.

The point of view that has stimulated the present study is that:

1) While the information provided by the temporal pattern of zero-axis crossings is probably necessary, it is possible that there are specific specimens of speech for which it may not be sufficient, and

2) While the information provided by the dynamic pattern of amplitude variations is probably not necessary for intelligibility of many speech samples, it is possible that there are some specific speech sounds for which this information may be necessary.

<u>Applications of peak clipping in communication</u>

Thus far infinite peak clipping has been treated in terms of distortion. It may also be viewed as a means of

simplifying the speech wave for efficient transmission.

Speech encoding. Licklider and Pollack (1948) observe that infinite peak clipping can reduce speech to a bivariate code more efficiently than pulse-modulation procedures. Not only is the speech wave easily encoded by infinite peak clipping, but may be decoded by the human ear with no further decoding apparatus necessary, other than the transducer that would ordinarily be used in the transmission system. Southworth (1963) describes several speech digitizing techniques based upon infinite clipping; these include pulse-number modulation and delta modulation techniques.

Broadcast transmission. One of the earliest applications of peak clipping was in broadcast transmission. Premodulation clipping has been found to increase the efficiency of power use in braodcast transmission (Kryter, Licklider, and Stevens, 1947).

Protective limiting. Clipping has been found to be useful in protecting the ear from high-energy speech peaks (Pollack and Pickett, 1959). This was investigated in detail for applications in hearing aid design (Davis, et al., 1947).

In all of the above applications, the ultimate usefulness will depend to some extent upon how intelligible peak clipped speech remains in relation to the needs of the specific situation. It is already known that word intelligibility is satisfactory under peak clipping. To the extent that specific

communication situations demand fine discriminations, the intelligibility of peak clipped speech must be determined in greater detail.

Limitations of previous studies on peak clipping

Influence of test materials. Experimental results on intelligibility are highly dependent on the type of test materials used. Under identical acoustic conditions, an intelligibility score may vary by as much as ninety percent depending upon whether the test materials are digits or nonsense syllables (Miller, Heise, and Lichten, 1951). Differences in intelligibility have been associated with meaning, number of syllables, and, to a smaller extent, syllabic stress (Hirsh, Reynolds, and Joseph, 1954). Since the currently available information about intelligibility of speech subjected to infinite peak clipping is based on monosyllabic words, the intelligibility of other speech samples, such as individual phonemes, remains unknown until tested empirically.

Effect of contextual clues. An important factor is the information provided by the context of sounds in a word or syllable; this may affect the number of alternative responses available if a particular sound is unintelligible (Miller, 1951). If, for instance, a test word is /stræ_/, with the final consonant unintelligible to the observer, and the test consists of nonsense syllables, there are more than twenty possible final consonants available as responses. These include the

nonsense words: _stram_, _strat_, _strab_, _strad_, _stran_, _strak_, _strang_, _strag_, _stral_, _strav_, _straf_, _straz_, _strath_ (voiced and voiceless), _stras_, _strash_, _strach_. If, however, the test vocabulary consisted of real English words, then there is probably but one possible response: strap. Thus, if the final sound were rendered unintelligible by some distortion, on the real word test the observer's knowledge of the statistical probabilities of occurrence of certain sounds, combinations of sounds, and orders of sounds would probably enable him to correctly identify the word. If a distortion, such as infinite peak clipping, produced a systematic effect on the intelligibility of certain sounds, this could be obscured in PB-50 scores if word contexts led to correct identification of words in which the sounds occurred.

As J. D. Harris (1960) has observed:

> In order to uncover the contribution of each type of physical cue alone, it is not enough to eliminate it by some legerdemaine in the laboratory. When this is done . . . intelligibility is often unaffected. A false conclusion could in that case be reached . . . that the cue eliminated is of minor importance in speech communication. What is necessary is to eliminate progressively one, two, and more cues simultaneously.

Since contextual clues may contribute to the intelligibility of distorted speech, it is considered important in this study to control contextual clues while studying the effects of distortion on intelligibility of individual sounds.

Statement of the Problem

This study addresses itself to three questions:

1) Are individual speech sounds affected in intellibibility by infinite peak clipping?

2) If so, are these effects equal from one phoneme to another?

3) If individual speech sounds are not equally affected in intelligibility by infinite peak clipping, is there a systematic pattern in the way in which sounds are affected?

CHAPTER II

PROCEDURE

Stimulus materials

Selection of stimuli. For this study twelve consonants /p b t d k g θ ð f v s z/ were chosen which represent a variety of positions and manners of articulation, as well as acoustic effects. In particular, the selected consonants include examples of fricatives and plosives, of voiced-voiceless cognate pairs, and of lingua-alveolar, lingua-dental, bilabial, and velar placements. They also represent the following distinctive features: contrasts of strident-mellow, tense-lax, continuous-discontinuant, grave-acute, and compact-diffuse.

Obviously some of these consonants, namely the plosives, cannot be spoken without an adjacent vowel. Moreover, it is well known that the acoustic characteristics of consonants can be altered by the vowel context in which they are produced. Consequently, in order to counterbalance the vowel environment, the selected consonants were embedded in a series of VCV syllables. Specifically, a corpus of ninety-six nonsense Vowel-Consonant-Vowel words was designed, drawing upon

the twelve consonants in combination with four stressed vowels and the neutral schwa sound. The counterbalancing was accomplished by selecting stressed vowels /i æ ɑ u/, which represent extremes of position on a vowel quadrilateral, and combining them with each of the consonants in both the pre-vocalic and post-vocalic positions. By adding the schwa to each of these CV and VC combinations in the initial or final position, respectively, VCV syllables were formed which correspond to Trochaic or Iambic stress patterns. These may be schematized as either stressed vowel + consonant + schwa, or schwa + consonant + stressed vowel. Thus for any replication of the entire corpus there are eight utterances of each consonant embedded in different vocalic environments. The entire set of stimulus words is presented in Table 1.

Because there are individual differences in the production of speech sounds, it was deemed desirable to use three different talkers to produce samples of the syllables. The three adult male talkers were chosen from the Communication Sciences faculty. Each talker had extensive training and experience in phonetic transcription.

In preparation for recording spoken examples of the selected stimuli, a phonetic transcription was made of the ninety-six nonsense syllables. The transcriptions were replicated three times, once for each talker, and arranged in one complete random order. On the final copies of this

Table 1. Vocabulary of nonsense syllables used as stimuli.

isə	æsə	ɑsə	usə
əsi	əsæ	əsɑ	əsu
iθə	æθə	ɑθə	uθə
əθi	əθæ	əθɑ	əθu
ifə	æfə	ɑfə	ufə
əfi	əfæ	əfɑ	əfu
izə	æzə	ɑvə	uzə
əzi	əzæ	əvɑ	əzu
iðə	æðə	ɑðə	uðə
əði	əðæ	əðɑ	əðu
ivə	ævə	ɑvə	uvə
əvi	əvæ	əvɑ	əvu
ipə	æpə	ɑpə	upə
əpi	əpæ	əpɑ	əpu
itə	ætə	ɑtə	utə
əti	ətæ	ətɑ	ətu
ikə	ækə	ɑkə	ukə
əki	əkæ	əkɑ	əku
ibə	æbə	ɑbə	ubə
əbi	əbæ	əbɑ	əbu
idə	ædə	ɑdə	udə
ədi	ədæ	ədɑ	ədu
igə	ægə	ɑgə	ugə
əgi	əgæ	əga	əgu

list of 288 phonetically transcribed syllables, individual items were identified by talker so that the recordings could be made in a single session.

Recording the stimulus tape. For two reasons, the spoken examples of stimuli were tape recorded during a single session. First, if recordings were made separately a subsequent splicing or rerecording step would have been necessary to randomize the stimuli produced by different talkers. However, by recording all the talkers in a single session, this randomization was accomplished prior to the recordings. Secondly, since the talkers were instructed to correct any errors in the production of the stimuli, they and the experimenter provided an initial check of each other's articulation.

In the recording session the talkers were seated in a series 1200 Industrial Acoustic Corporation room approximately two feet from the Altec M-20 microphone system, which was coupled to an Ampex model 351-C full-track tape recorder located outside the sound-treated room. Each talker-listener had a complete list of the stimuli indicating which talker was to read each of the words. Talkers were instructed to use the carrier phrase "now" for each VCV word with no intervening pause. By using such a phrase an opportunity for the talker to reach a stable vocal level was provided before his production of the VCV syllable. The word "now" was chosen, rather than a more conventional "say the word" or "write the

word," because it does not include any of the sounds used as stimuli. This avoided the possibility of a constant standard of reference for one or more of the test stimuli. The experimenter provided a flash of light as a cue for the beginning of each word in order to space stimuli at four-second intervals.

Following the recording session, the tape was edited to remove errors and irrelevant materials and to insert an identifying number before each group of ten stimuli. Since one of the talkers tended to use a slightly lower vocal level than the others, the gain was adjusted during the rerecording process to equalize the levels of the speech peaks on the carrier phrase.

<u>Observations</u>

<u>Observers</u>. The criteria for selection of observers were: 1) training in phonetics, and 2) normal speech and hearing. Hearing sensitivity within normal limits was determined by means of individual screening tests at 250, 500, 1000, 2000, 4000, and 6000 Hz. This testing was accomplished with a Beltone model 10 A portable audiometer, set at a screening level of 15 dB relative to the 1964 I.S.O. standard.

Ten observers who met the criteria were recruited from among the students and faculty associated with the Communication Sciences program at the University of Florida. This included eight male and two female observers whose ages ranged

from twenty to fifty years. For each of the observers, the experimental session was the first exposure to speech subjected to infinite peak clipping.

Playback system. The complete system as used in the experimental trials is depicted schematically in Figure 2. The stimulus tape was played from an Ampex 351-C tape recorder into one channel of a Marantz model 7 preamplifier with controls set in the normal flat frequency response position. The output of this channel was led to a double-pole double-throw selector control which either connected or bypassed the clipper. The clipping unit, consisting of three cascaded printed-circuit amplifier modules, operated beyond their linear range, is described in detail in Appendix A. The "pass-clip" switching control output led to a Hewlett-Packard 350-D decade attenuator set. The signal was then reamplified through the second channel of the preamplifier connected to a Marantz model 8-B power amplifier. The output of the power amplifier drove an AR-3 acoustic suspension loudspeaker system located in the I.A.C. room. A dual-beam Tektronix type 564 oscilloscope was used to constantly monitor the speech waveforms at the input of the clipping unit and at the output of the power amplifier. Accordingly, both undistorted and clipped waveforms could be observed simultaneously.

Playback level. Stimuli were presented to observers at 70 dB Sound Pressure Level. In order to achieve the desired

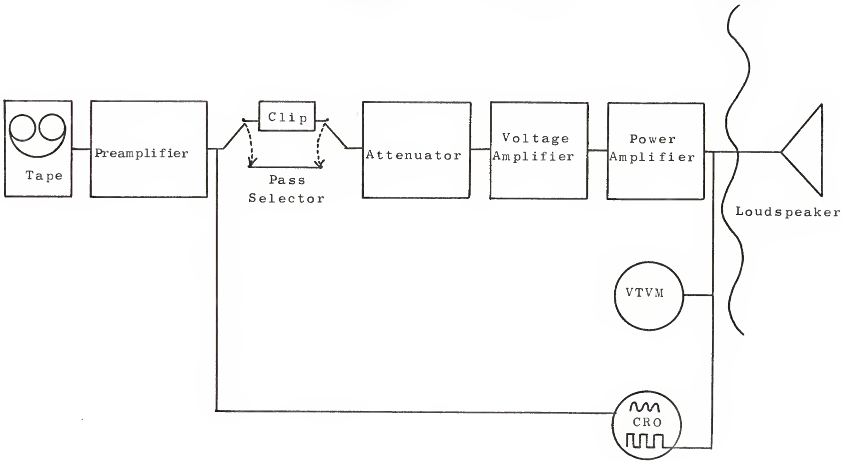


Figure 2. Block diagram of equipment.

level consistently for all experimental trials, a preliminary calibration was performed. A General Radio sound level meter, located at a position in the I.A.C. room, where observers would sit during experimental trials, was used to determine sound pressure levels. The RMS voltage reading at the input to the loudspeaker which produced a 70 dB Sound Pressure Level at the observer position was noted and used as reference setting during experimental trials. The signal used to set voltage readings was a 1000 Hz calibration tone. The calibration tone was recorded on the test tape at a V.U. level that corresponded to the average V.U. of the talkers' speech peaks on the tape; the signal source for the recording was a General Radio model 1304-B beat frequency audio generator. RMS voltages at the transducer input were determined with a Ballantine model 321-C A.C. Vacuum Tube Voltmeter. Sound Pressure Levels were read on the B scale of the sound-level meter.

System calibration. In order to determine whether the present experimental apparatus was comparable with that of previous studies, a system calibration was performed. This took the form of a pilot study using a speech sample similar to that used in the Licklider and Pollack (1948) study. Details are presented in Appendix B. The apparatus was found to be comparable.

Conducting experimental trials. Observers were seated

in one of three seats placed at a distance of two and one-half feet from the loudspeaker. Half of the observers heard the test list under the "clip" treatment first and "pass" (undistorted) second; the other five observers heard the lists with treatment conditions in the reverse order.

The following recorded instructions were presented to the observers before the VCV list under each treatment condition:

> Transcribe only the consonant in the word. Do not transcribe the word "now" which comes before each word. For example, if you hear "Now ɛʃə," write the symbol for /ʃ/. If you hear "Now əmɔ," write the symbol for /m/. To help you keep your place, after each group of ten words the number of the following word will be given.

Just prior to the instructions for the "clip" condition, whether it came first or second in order, certain recorded materials were presented to provide some familiarization with the sound of clipped speech. These materials included one unfamiliar passage--a paragraph from an out-dated news magazine, and two passages which were presumed to be very familiar: the pledge of allegiance to the flag, and digits from one to twenty.

A brief rest period was provided between treatment conditions. At this time the voltage to the loudspeaker was rechecked and the tape was advanced to the position for the following treatment. The listening session required about

one hour to complete both treatment conditions.

Data reduction

Analysis of variance. The primary analysis of results used a design with four factors: Consonant (C), Observer (O), Talker (T), and Distortion (D). Factor C consisted of the twelve consonants used in this study, Factor O represented ten observers and Factor T represented three talkers. The two treatments in Factor D were the distorted, or "clip" condition and the undistorted, or "pass" condition. The latter was included as the experimental control for the effects of clipping. The design utilized a mixed model in which the Consonant and Treatment factors were considered fixed effects and the Observer and Talker factors were considered random effects. The structural design is depicted in Figure 3. The number of correct responses (from zero to eight) made by a single observer to a talker's productions of a given consonant was entered in the individual cell. For purposes of this analysis, differences among vowels and stress patterns in syllables were ignored; the syllable was viewed simply as the vehicle for providing a means of counterbalancing a variety of coarticulation effects among the test stimuli, and as a method of limiting the number of response alternatives for the total VCV vocabulary in order to control the contextual variable.

Individual comparisons. A factorial design makes

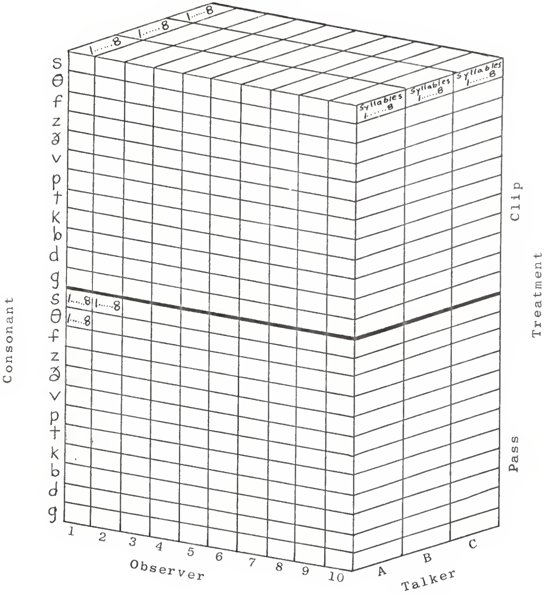


Figure 3. Structural scheme of factorial design.

possible the analysis of two or more experimental variables simultaneously, both of their individual effects and of their interactions with one another. In an analysis of variance, the $\underline{F}$ ratios obtained are evaluated with reference to the theoretical distribution of $\underline{F}$ in order to test the significance of differences among the treatment means.

It is often deemed desirable to investigate the sources contributing to the significance of a given variable. This may be accomplished by making specific comparisons of the significances of differences among means of the levels or categories of a variable. For instance, in the present study the information that the Consonant factor had statistical significance would not indicate which consonants differed from one another, but comparisons among the individual consonants or groups of consonants could provide this information.

For enumerative data such as those collected in this study the chi-square test provides an appropriate test of significance. The present data represent the number of correct responses that were actually obtained compared with the number of correct responses possible. According to Snedecor (1956) the basic chi-square formula may be verbalized as: "Chi-square is the sum of such ratios as (deviation squared)/(hypothetical number)." The factorial chi-square, developed by Brandt and Snedecor (Brandt, 1949), represents an alternative method of calculation of chi-square that is easily

adaptable to use in a factorial design. In this method the sums of squares are multiplied by a chi-square factor to determine the value of chi-square associated with a given comparison. The chi-square factor represents $(e)^2/(o)(e-o)$, where $\underline{e}$ equals the total possible sum of scores and $\underline{o}$ equals the obtained sum of scores. The sum of squares associated with the comparison is $(d)^2/(f)(s_1^{\,2}+s_2^{\,2})$, where $\underline{d}$ equals the difference between sums of scores of the quantities being compared, $\underline{f}$ is the frequency of each item being compared, and $\underline{s}_1$ and $\underline{s}_2$ are the number of items comprising each half of the comparison. For example, if one sound is compared with another, $(s_1^{\,2}+s_2^{\,2})$ would equal twice one squared, or two; if six sounds were compared with six other sounds, $(s_1^{\,2} + s_2^{\,2})$ would equal twice six squared, or seventy-two. The obtained value of chi-square is tested for the number of degrees of freedom appropriate to the effect or interaction for which the comparison is made.

This method was used to make voiced-voiceless comparisons, fricative-plosive comparisons, and individual consonant comparisons. It was also used to test the effect of order of presentation of treatments.

Stimulus-response matrices. Information about types of error responses was displayed in confusion matrices which were compiled for each observer (see Appendix C), similar to those described by Miller and Nicely (1955). These matrices preserved

specific responses made by observers to each talker's stimuli.

CHAPTER III

RESULTS AND DISCUSSION

Analysis of Variance

A summary of the results of the analysis of variance is presented in Table 2.

Main effects

Two main effects were found significant. First, the Distortion factor was significant, indicating that clipping does produce a decrement in consonant intelligibility. The total percent of correct responses under the "pass" condition was 94.6%; the corresponding score for clipped speech was 76.3%. This gross effect of clipping on all stimuli together does provide an affirmative answer to the experimenter's first question, are individual speech sounds affected by infinite peak clipping?

The Consonant main effect was also significant. This may be due to an inherent difference in consonant intelligibility, whether clipped or undistorted. It may be noted in Table 3 that for both /θ/ and /ð/, intelligibility scores were lower than for other consonants under both conditions.

Table 2. Summary of analysis of variance.

Source	d.f.	Mean Square	F_1	Significance[2]
Consonant (C)	11	92.45	9.44	**
Observer (O)	9	3.18	0.33	N.S.
Talker (T)	2	0.05	0.01	N.S.
Distortion (D)	1	387.20	39.55	**
CxO	99	13.07	1.29	*
CxT	22	52.05	5.14	**
CxD	11	142.71	14.08	**
OxT	18	2.23	0.22	N.S.
OxD	9	48.99	4.83	**
TxD	2	200.53	19.79	**
CxOxT	198	7.81	0.63	N.S.
CxOxD	99	20.57	1.65	**
CxTxD	22	80.19	6.44	**
OxTxD	18	26.69	2.14	**
CxOxTxD	198	12.45		
pooled residual_a				
OxT				
CxOxT				
CxOxTxD	414	9.79		
pooled residual_b				
CxOxT				
CxOxTxD	396	10.13		

[1] error term for main effects was pooled residual_a
error term for two-way interactions was pooled residual_b
error term for three way interaction was CxOxTxD

[2] .01 level of significance **
.05 level of significance *
not significant N.S.

Table 3. Difference between treatments for each consonant.

Consonant	Percent Correct Pass	Percent Correct Clip	Chi-square
s	98.8	94.6	1.67
θ	85.8	38.8	213.56
f	89.6	70.8	33.87
z	99.2	83.3	25.44
ð	65.8	27.5	141.56
v	97.9	74.6	52.45
p	100.0	87.5	15.05
t	99.6	91.3	6.69
k	99.6	83.8	24.15
b	99.2	84.2	21.68
d	99.6	87.9	13.11
g	100.0	91.3	7.38

With 11 df, chi-square needed for significance at .05- 19.7
.01- 24.7

The main effects for Talker and for Observer were not found to be significant. This was reassuring, since both of these factors were potential sources of considerable experimental error. Since over-all differences among talkers and observers were not statistically significant, many of the scores to be reported are based upon the pooled talker and observer scores.

Interactions

Despite the lack of significance of the Talker and Observer main effects, there were several significant interactions involving talkers and observers. Visual inspection of the data indicated that the Consonant x Talker interaction may have been due to a higher score of correct responses to /θ/ as spoken by talker A than by the other talkers, and a higher score on /ð/ as spoken by talker C. Furthermore, the Talker x Distortion interaction showed significance since talker C yielded a higher score than talkers A and B for the "pass" condition but a lower score than the others under clipping. These interactions involving talkers tend to confirm the advisability of using more than one talker in a study such as this. Divergent over-all scores under clipping between two observers seemed to be the basis of the significant Observer x Distortion interaction.

Of major importance in this study is the Consonant x Distortion interaction. The significance of this

interaction provides an answer to the experimenter's second question; are these effects equal? The effect of clipping on individual sounds is not equal.

Individual Comparisons

Specific consonants

The difference between distortion treatments for each consonant is presented in Table 3, in which the percentage of correct responses is compared for each consonant. For almost half of the consonants that were included as stimuli, the intelligibility under clipping was not significantly different than for undistorted speech. The sounds that were affected little were /s g t d p/. Significantly different at the .01 level were /ð θ v f/. Of the consonants that were affected by clipping, the decrement in percent of correct responses was as great as 47%. These relationships may be visualized easily in Figure 4.

Considering each treatment separately, the difference from one consonant to another is shown in Table 4. This table is arranged in two matrices; the upper matrix represents the "pass" condition and the lower represents the "clip" condition. In each matrix, the presence of an asterisk at the intersect of a column and row indicates a statistically significant difference between the two corresponding sounds in the margins. A blank space indicates that for the treatment under consideration, the difference between the two sounds

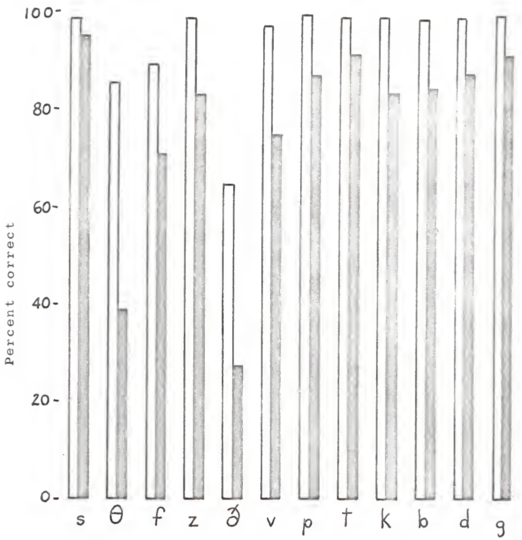


Figure 4. Intelligibility scores of consonants.

A. Open bar indicates undistorted condition

B. Closed bar indicates infinite peak clipped condition

Table 4. Matrix of significant differences[1] between sounds under each distortion treatment.

	Pass											
	s	t	g	d	p	b	k	z	v	f	θ	ð
s (95.0%)												*
t (91.3%)												*
g (91.3%)												*
d (87.9%)												*
p (87.5%)												*
b (84.2%)												*
k (83.8%)												*
z (83.3%)												*
v (74.5%)	*	*	*									*
f (70.8%)	*	*	*	*								*
θ (38.8%)	*	*	*	*	*	*	*	*	*	*		*
ð (27.5%)	*	*	*	*	*	*	*	*	*	*		
	Clip											

Sounds are arranged in order of decreasing intelligibility under clipping.

[1]significance level .01

(with respect to the number of stimuli correctly identified) is not statistically significant at the .01 level. Under the "pass" condition, /ð/ is statistically different from all of the other sounds. Under the "clip" condition, there are four consonants that are uniformly less intelligible than the others; these are /ð θ f v/. Reference to Table 4 makes it possible to ascertain the significance of the difference in intelligibility, under either treatment, between any two of the test stimuli.

Results of the present study seem to support the observations made by Licklider and Pollack that a speech waveform can retain intelligibility even when stripped of relative amplitude clues so that only the temporal pattern of zero-axis crossings is retained. In the present study, data are also presented which show that the extent to which this is true depends upon the speech sample.[1] With contextual clues controlled, the intelligibility of individual consonant sounds in VCV syllables ranged from below thirty percent to scores above ninety percent, for observers' first exposure to speech distorted by infinite peak clipping. The differential effect

---

[1]A pilot study was conducted as a calibration procedure to determine the equivalence of the present experimental situation to that of the Licklider and Pollack study. It was concluded that the two were comparable. See Appendix B for details.

of clipping on individual consonants is emphasized by noting the scores for the sounds that were the most affected and those that were the least affected by clipping. Scores for /θ/ and /ð/ were 38.8% and 27.5%; compared with the scores obtained without distortion, these represent decrements of 47.0% and 38.3% respectively. The score of 95.0% for /s/ represents a decrement of only 3.6%. which was not statistically significant.

Voiced-voiceless grouping

The value of chi-square associated with the grouped scores of /s θ f p t k/ and grouped scores of /z ð v b d g/ under clipping was 0.36; the value necessary for significance at the .05 level was 19.70. There were no significant differences under clipping between pairs of voiced-voiceless cognate sounds, as shown in Table 5. Thus, clipping produced the same degree of effect on intelligibility in the voiced and voiceless sounds. Since these groups are primarily differentiated on the basis of the presence or absence of low-frequency periodic excitation, it is to be expected that clipping would not affect recognition of voicing.

Fricative-plosive grouping

In comparing consonants grouped according to the fricative-plosive classification, a value of chi-square of 46.14 was obtained; the value necessary for significance at the .01 level was 24.70. It is apparent that the fricatives

Table 5. Differences between voiced-voiceless cognate pairs subjected to infinite peak clipping.

Consonant Pair	Chi-square
S - z	14.21
θ - ð	12.19
f - v	1.35
p - b	1.07
t - d	1.07
k - g	5.43

With 11 df, value of chi-square needed for significance at
.05 - 19.7
.01 - 24.7

were different, as a group, from the plosives. Intelligibility of the fricatives was considerably lower. However, this must be qualified; two of the fricatives, /s/ and /z/ were relatively unaffected by clipping, while the other four, /θ ð v f/ were severely degraded in intelligibility. Consequently, it appears that the distinguishing characteristics of fricatives are obscured more than those of plosives, and furthermore, that the characteristics of the dental plosives are obscured more than the alveolar plosives. This provides an answer to the third question, is there a systematic way in which sounds are affected by peak clipping? The following reasons are offered in partial explanation:

1) Cavity resonance. The four severely affected fricatives are formed at the teeth, so far forward in the mouth that they are relatively free of oral cavity resonance effects. The friction component of these sounds consists only of the turbulence of the breathstream between tongue and teeth or between lip and teeth, radiating into the air (plus voicing for /ð/ and /v/). It is evident that without oral cavity resonance there is little basis for distinguishing the lingua-dental sounds from the labio-dental sounds.

2) Friction and vocalic components of fricatives. An experiment performed at the Haskins Laboratories seems especially relevant to this discussion. Harris (1958) separated the friction components of fricative sounds in VC

syllables from the vocalic portions and combined the friction component of one fricative with the vocalic portion associated with a different fricative. She found that identification of /s/ and /z/ could be made by observers entirely on the basis of the friction component, but that identification of /f v θ ð/ depended on the combination of both friction and vocalic portions. Since in the present study /s/ and /z/ retained such high intelligibility, it might be inferred that the removal of relative amplitude clues did not affect the transmission of information inherent in the friction portion of fricatives, but that it did affect the vocalic portion.

3) Balance of noise. The /θ/ and /ð/ are characterized by a low intensity noise, while the /s/ and /z/ are considered to have a high intensity noise component (Jacobson, Fant, and Halle, 1952). Since one characteristic of infinite peak clipping is to equalize amplitudes to a uniformly high level, it is possible that clipping might do more violence to the naturally lower noise levels of /θ/ and ð/ than to the normally higher noise levels of /s/ and /z/.

4) Durational clues. Still another explanation can be made on the basis of the clue of duration. While traditionally the fricatives are considered longer in duration than plosives, Miller and Nicely (1955) included /s/ /z/ and the lateral sibilant fricatives in one class, representing the consonants of longest duration, but the dental fricatives were placed in

the same category of duration as the plosives. Since temporal characteristics would be unchanged by clipping, this would provide an additional means of identifying the /s/ and /z/. It would also help to explain certain errors that occurred, in which /p t k/ were the responses to /f/ and to /v/.

5) Affrication in plosives. A factor related to confusions between fricatives and plosives may be the affrication that typically follows plosives in American speech. This affrication, when amplified by peak clipping, might tend to make the clipped plosives perceptually resemble the fricatives.

Order of treatment

The possibility of an order effect based on the procedure in which five observers received the "pass" condition first and the other five received the "clip" condition first was investigated. As Table 6 shows, the order of presentation did not yield a significant difference for either the "pass" or the "clip" condition.

Error Responses

Confusion matrices were prepared which combined results for all talkers and for all observers. Figures 5 and 6 show types of responses actually made to the stimuli and their frequency of occurrence. The individual scores shown in these figures are based on ten observers' pooled responses to the eight syllables spoken by three talkers for each consonant--a total of 240 stimuli per consonant.

Table 6. Effect of order of treatment for each treatment.

Treatment	Order: Pass First (Sum of Scores)	Order: Clip First (Sum of Scores)	Difference	Chi-square
Pass	1374	1350	24	1.60
Clip	1135	1061	74	15.26

With 9 df, value of chi-square needed for significance at
.05 - 16.9
.01 - 21.7

Percent response on pass treatment

Stimuli	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	98.75	1.25																				
θ		85.83	12.08		0.83		0.83			0.41												
f		9.16	89.58		1.25																	
z	0.41			99.16	0.41																	
ð		6.66	0.41		65.83	25.83					0.83	0.41										
v		0.41			1.66	97.91																
p							100															
t								99.58	0.41													
k								0.41	99.58													
b										99.16	0.41		0.41									
d										0.41	99.58											
g												100										

Figure 5. Confusion matrix of responses to undistorted consonants.

Percent response on clip treatment

Stimuli	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	94.58	2.91	1.66	0.41				0.41														
θ	0.83	38.75	27.91		1.25		8.75	9.16	9.58	0.83			0.83							1.25		0.83
f		7.08	70.83				14.16	0.41	5.00	0.41										0.83	0.83	0.41
z	0.83			83.33	8.75	3.33					0.83	1.25		0.41	0.41	0.41	0.41					
ð		4.16	0.41	2.08	27.50	43.75				0.83	1.25	2.91		6.25	5.83		2.50	2.08	0.41			
v		0.41	0.41	0.83	9.16	74.58	0.83			2.08			0.41	2.91	5.42	0.41	2.08	0.41				
p					0.41		87.50	2.91	7.50			0.41		1.25								
t							4.16	91.25	4.58													
k							2.91	12.50	83.75				0.41									
b					0.41	5.83	0.41			84.16	2.50	3.75	0.41	0.83			1.66					
d					0.83	0.83		0.41		3.75	87.91	5.83		0.41								
g						1.25		0.41	0.41	0.41	3.75	91.25						2.08	0.41			

Figure 6. Confusion matrix of responses to consonants distorted by infinite peak clipping.

Generally, under the "pass" condition most response errors could be classified in the same voicing and manner of release category as their respective stimuli. However, under the "clip" condition the types of error were divergent. There were very few instances of voiced-voiceless confusions, but there were many confusions on the plosive-fricative dimension. These tended to be unidirectional; that is, fricatives, especially /θ/ and /ð/, were frequently identified as plosives, but plosives were rarely heard as fricatives. There were other interesting examples of confusions that were not reciprocal: twenty-eight percent of the responses to /θ/ were /f/, but only seven percent of the responses to /f/ were /θ/. Similarly, forty-three percent of the responses to /ð/ were /v/, but only nine percent of the responses to /v/ were /ð/. Under the "pass" condition the confusions between /θ/ and /f/ were approximately equal, but confusions between /ð/ and /v/ exhibited the same sort of asymmetry as under clipping.

There were many more types of error responses used by observers under the "clip" than the "pass" condition. This was more noticeable for the dental fricatives than for other sounds.

CHAPTER IV

SUMMARY AND CONCLUSIONS

Summary

An experiment was performed to determine whether individual consonant sounds are affected differentially by infinite peak clipping. The consonants, representing plosives and fricatives, voiced and voiceless, and five articulatory positions, were placed in Vowel-Consonant-Vowel nonsense syllables. These stimuli were spoken by three trained talkers following a predetermined random word order and were tape recorded. Ten adult observers, trained in phonetics, transcribed the consonants in the stimulus words under two conditions: one in which the speech was passed through the system undistorted, and one in which the speech was distorted by infinite peak clipping.

The data were subjected to a factorial analysis of variance to appraise the statistical significance of the results. The main effect for Distortion treatment was significant as was the interaction between Distortion and Consonant. This indicated that not only did infinite peak

clipping affect the intelligibility of speech, but it does so differently for different consonant sounds. The main effects reflecting Talker and Observer variance were not significant, indicating a lack of systematic error from these sources.

The dental fricatives were the most severely affected by clipping. The plosives and /s/ and /z/ were the least affected by clipping. Intelligibility scores for clipped consonants ranged from 27.5% to 94.6%; the range for the undistorted condition was from 65.8% to 100%. Sounds that exhibited the poorest intelligibility when not distorted also showed the lowest intelligibility under clipping. Voiced and voiceless sounds were equally affected by clipping.

Conclusions

Infinite peak clipping affects the intelligibility of individual speech phonemes.

The effect of infinite peak clipping on the intelligibility of individual consonant sounds is different from one sound to another.

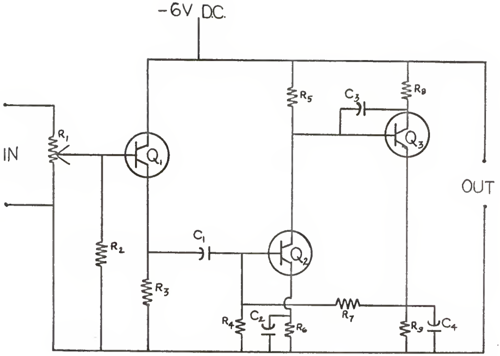
There is a systematic pattern in the relative effects of infinite peak clipping on intelligibility of individual sounds. Dental fricatives show the largest decrement under clipping; plosives and sibilant fricatives are the least affected of the sounds sampled. Intelligibility of voiced and voiceless sounds are equally affected by infinite peak clipping.

APPENDIX A

DETAILS OF THE CLIPPING APPARATUS

Three resistance-capacitance-coupled printed-circuit amplifier modules were used in cascade. Clipping was achieved by overdriving the amplifiers between saturation and cutoff, as described by Littwin (1964). Each module consisted of one emitter-follower stage driving two stages of common-emitter amplification. A schematic of the amplifier modules is shown in Figure 7.

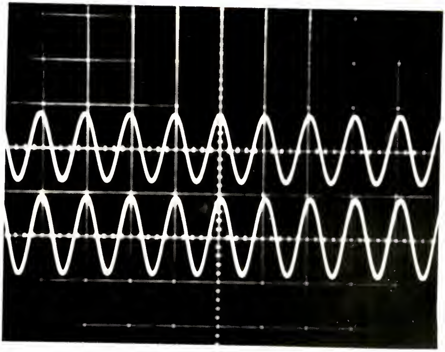
Undistorted operation of each module separately and of the three modules cascaded, when operated within their linear range, was demonstrated by observing the input and output waveforms on a Tektronix type 564 dual-beam storage oscilloscope. These waveforms were photographed using a C-12 oscilloscope camera with Polaroid pack. In Figure 8 the upper traces show a 1000 Hz tone at the output of a Hewlett-Packard model 201-C oscillator, and the lower traces show the waveforms at the output of successive cascaded modules. In Figure 8A the unit is shown operating within its linear range; in B it is operating as a peak clipper. The action of the clipper on samples of speech is shown in Figure 9. Photographs A and



R_1	3,000 ohms		C_1	10. microfarads
R_2	47,000 ohms		C_2	100. microfarads
R_3	10,000 ohms		C_3	.001 microfarads
R_4	4,700 ohms		C_4	100. microfarads
R_5	4,700 ohms			
R_6	680 ohms		Q_1	Low Beta p-n-p
R_7	3,300 ohms		Q_2	Low Beta p-n-p
R_8	2,100 ohms		Q_3	High Beta p-n-p
R_9	1,500 ohms			

Figure 7. Schematic of clipper amplifier module.

A



B

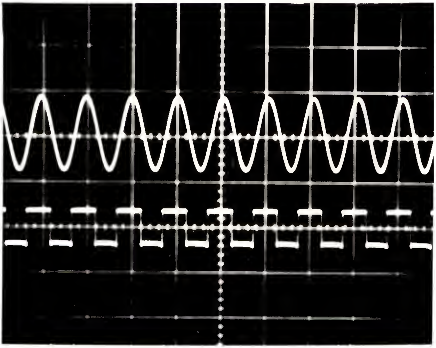
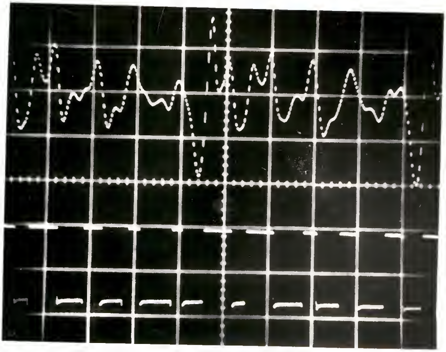


Figure 8. Output waveforms of modules
A. Linear operation
B. Peak clipping

A



B

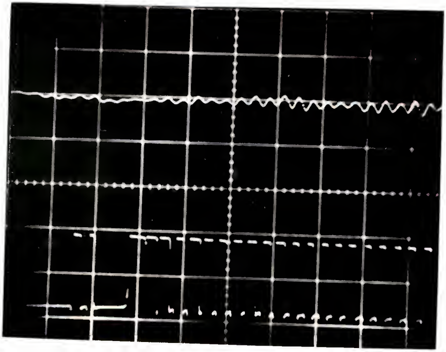


Figure 9. Action of clipper on speech waveforms.
A. Vowel
B. Consonant

B show waveforms at two different moments during utterance of the word "mash," corresponding to the vowel and the final consonant. The upper traces show the speech signal at the tape playback output of an Ampex model 354 tape recorder; the lower traces show the same samples monitored at the output of the clipper.

The frequency response of each module and of the three in tandem was found to be 60 to 12,000 Hz $\pm$ 3 dB. This was determined by measuring the root mean square output voltage of the clipper, with constant input, using pure tones generated by a General Radio model 1304-B beat frequency audio oscillator. Voltages were measured with a Ballantine model 321 A.C. Vacuum Tube Voltmeter.

APPENDIX B

SYSTEM CALIBRATION PILOT STUDY

Purpose

To determine if the present experimental situation was comparable to that used by previous investigators, the traditional procedures were replicated in a pilot study. In particular, the intelligibility of undistorted and clipped PB words was compared using materials described by Licklider and Pollack (1948).

Stimulus materials

Lists one and two of the Harvard PB-50 words (Egan, 1948) were arranged in random word order. As in the Licklider and Pollack (1948) study, one adult male talker who had training in phonetics read both lists. The words were tape recorded in a series 1200 Industrial Acoustics Corporation room using an Altec M-20 condenser microphone coupled to an Ampex model 350-C full-track tape recorder.

The stimulus words, each in the carrier phrase "Write the word _____," were spoken as one continuous sentence. During the recording, the talker monitored the deflections of a VU meter associated with this phrase in order to maintain a

uniform level of vocal intensity among the words. The words were separated by an interval of five seconds, and before each ten words a number identifying the following stimulus was given, using the phrase, "Next is number _____."

Apparatus

The peak clipping unit consisted of three cascaded printed-circuit amplifier modules operated beyond their linear range. Details of this instrument are given in Appendix A. A high quality audio playback system which is described in Chapter II was used for presentation of stimulus materials.

Subjects

Ten observers, three males and seven females, were selected from the staff of the Communication Sciences Laboratory. The observers ranged in age from nineteen to thirty-six years and demonstrated normal hearing. Hearing was screened at 15 dB relative to the I.S.O. standard at test frequencies of 250, 500, 1000, 2000, 4000, and 6000 Hz, using a Beltone model 10 A portable audiometer.

Procedure

Experimental trials were conducted in a series 1200 Industrial Acoustics Corporation auditory test room. Observers were seated in one of three chairs positioned at two and one-half feet from the loudspeaker. Instructions were given to write the words that they heard, using response forms that were provided with spaces for the fifty words of a PB word list.

In order to avoid any systematic effect associated with the difficulty of one PB list relative to another, the distortion treatments and word lists were counterbalanced. For the "pass" treatment (without peak clipping), five observers heard List One, and the other five observers heard List Two. For the "clip" treatment, each observer heard the list that was not presented to him under the "pass" treatment.

Results

Results are shown in Table 7. The mean intelligibility scores for the ten observers for the "pass" treatment was 95.8%; the scores for the "clip" treatment were 68.8%. In the Licklider and Pollack study (1948) the mean score for five observers in the first test session was approximately 98% without distortion, and 64% with infinite peak clipping.

Conclusion

With PB-50 lists as stimulus material, the apparatus and procedure used in the present experiment duplicated results obtained by Licklider and Pollack (1948). Intelligibility scores obtained without distortion and with infinite peak clipping were of the same order of magnitude as those reported in the literature.

Table 7. Intelligibility scores for PB lists without distortion and with infinite peak clipping.

Observers	No Distortion		Clipping	
Group 1	List 1	95.2%	List 2	65.6%
Group 2	List 2	96.4%	List 1	72.0%
All		95.8%		68.8%

APPENDIX C

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	7	1																				
	B	8																					
	C	8																					
θ	A		7					1															
	B		5	3																			
	C		7	1																			
f	A		1	7																			
	B		2	6																			
	C			8																			
z	A				8																		
	B				8																		
	C				8																		
ð	A					5	3																
	B					7	1																
	C					8																	
v	A						8																
	B					1	7																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 10. Matrix of responses by Observer 1 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		5						2	1													
	B		4	2																	2		
	C		4					1	1	1											1		
f	A			4				2		2													
	B		1	6				1															
	C		3	3																	2		
z	A				7	1																	
	B				8																		
	C				8																		
ð	A					4	1				1							2					
	B					1	3								1	1		1		1			
	C					3	1								4								
v	A			1			3				2					1		1					
	B					2	2								1	1		1	1				
	C					4	1								2	1							
p	A							7		1													
	B							2	3	2					1								
	C							6	2														
t	A								8														
	B								8														
	C							1	7														
k	A								2	6													
	B								3	5													
	C								2	6													
b	A										7	1											
	B										6							2					
	C										5				2			1					
d	A											8											
	B											8											
	C										2	6											
g	A												7						1				
	B												6						1	1			
	C										1	1	6										

Figure 11. Matrix of responses by Observer 1 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	3																					
	B	8																					
	C	7	1																				
θ	A		8																				
	B		7	1																			
	C		7	1																			
f	A		3	5																			
	B		1	6		1																	
	C			8																			
z	A				7	1																	
	B				8																		
	C				8																		
ð	A		1			6	1																
	B		1			6						1											
	C					8																	
v	A		1				7																
	B					2	6																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								7	1													
k	A									8													
	B									8													
	C									8													
b	A										7	1											
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 12. Matrix of responses by Observer 2 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	7			1																		
	B	8																					
	C	8																					
θ	A		5						1	2													
	B		2	4					1	1													
	C		3	2		1		2															
f	A		1	2				4		1													
	B			7																			
	C		1	7																		1	
z	A				8																		
	B				8																		
	C				7	1																	
ð	A		1			2	5																
	B				1	2	4											1					
	C					3	3						1			1							
v	A						8																
	B					1	7																
	C					2	6																
p	A							7		1													
	B							6		2													
	C							7	1														
t	A								7	1													
	B								8														
	C							1	7														
k	A								1	6				1									
	B								1	7													
	C								2	6													
b	A										8												
	B										7	1											
	C										7	1											
d	A											8											
	B											8											
	C										1	7											
g	A												8										
	B												8										
	C												8										

Figure 13. Matrix of responses by Observer 2 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

		s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	7	1																				
	B	8																					
	C	8																					
θ	A		8																				
	B		5	3																			
	C		8																				
f	A		1	7																			
	B			8																			
	C		1	7																			
z	A				8																		
	B				8																		
	C				8																		
ð	A		3			2	3																
	B		1			3	4																
	C		4			1	3																
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 14. Matrix of responses by Observer 3 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A	1		4					1	1				1									
	B		2	5					1														
	C		2	4				2															
f	A			5				2		1													
	B			7				1															
	C			8																			
z	A				7								1										
	B				8																		
	C				7											1							
ð	A						6					1						1					
	B		1		2		3									1			1				
	C						2						1		2	3							
v	A						5	1			1					1							
	B				1		5							1		1							
	C						4								2	2							
p	A							8															
	B							7		1													
	C							8															
t	A								8														
	B								8														
	C								8														
k	A							1	1	6													
	B							1	1	6													
	C								2	6													
b	A										8												
	B										7		1										
	C										7		1										
d	A											8											
	B											8											
	C											7	1										
g	A											1	7										
	B												8										
	C												8										

Figure 15. Matrix of responses by Observer 3 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli		s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		8																				
	B		6	1		1																	
	C		7	1																			
f	A		1	6		1																	
	B			7		1																	
	C			8																			
z	A				8																		
	B				8																		
	C				8																		
ð	A		1			3	4																
	B					3	4					1											
	C			1		5	2																
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 16. Matrix of responses by Observer 4 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		2	4						2													
	B		1	6						1													
	C		2	4						2													
f	A			7						1													
	B			8																			
	C			8																			
z	A	1			6								1										
	B				8																		
	C				7													1					
ð	A			1	1		5								1								
	B						6						1						1				
	C				1		5									2							
v	A						7									1							
	B						7								1								
	C						7									1							
p	A							6		2													
	B							5		2					1								
	C							5	1	2													
t	A								7	1													
	B								8														
	C								4	4													
k	A							1	1	6													
	B									8													
	C								1	7													
b	A										8												
	B										8												
	C						1				5		1					1					
d	A											7	1										
	B											8											
	C											6	2										
g	A												8										
	B												7						1				
	C												8										

Figure 17. Matrix of responses by Observer 4 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli		s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		8																				
	B		5	3																			
	C		6	2																			
f	A			8																			
	B			8																			
	C			8																			
z	A				8																		
	B				8																		
	C				8																		
ð	A					4	4																
	B					4	4																
	C					7	1																
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A								1	7													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 18. Matrix of responses by Observer 5 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

		s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	6		1					1														
	B	8																					
	C	6		2																			
θ	A			2				1	2	2	1												
	B		1	4		1			1														1
	C			3				2	1	1													1
f	A			3				3		2													
	B			7				1															
	C			5				1	1														1
z	A				4	2	2																
	B				8																		
	C				4	1	2						1										
ð	A						8																
	B				1		4						2		1								
	C						3				1				2	2							
v	A						5	1			1					1							
	B						7								1								
	C					1	6									1							
p	A							7					1										
	B							6		1					1								
	C							8															
t	A								7	1													
	B								8														
	C								8														
k	A							1		7													
	B									8													
	C									8													
b	A										8												
	B						1				5	1	1										
	C						1				7												
d	A								1			7											
	B											6	2										
	C											5	2		1								
g	A											1	7										
	B												7						1				
	C												8										

Figure 19. Matrix of responses by Observer 5 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		8																				
	B		7	1																			
	C		6	1							1												
f	A		2	6																			
	B		1	7																			
	C			8																			
z	A				8																		
	B				8																		
	C				8																		
ð	A					6	2																
	B					6	1						1										
	C					8																	
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 20. Matrix of responses by Observer 6 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		4					2		1				1									
	B		4	2		1			1														
	C		3	1				1	3														
f	A			3				3		1	1												
	B			6				2															
	C		1	6				1															
z	A				8																		
	B				8																		
	C				6	1						1											
ð	A					3	3					1			1								
	B					1	3						1					1	2				
	C					5	1								1	1							
v	A		1				6											1					
	B						7											1					
	C					1	5										1	1					
p	A							8															
	B							7		1													
	C							8															
t	A								8														
	B								8														
	C							1	7														
k	A									8													
	B								1	7													
	C									8													
b	A										8												
	B										5	1	1					1					
	C										7			1									
d	A											8											
	B											8											
	C										1	7											
g	A											1	7										
	B											1	7										
	C												8										

Figure 21. Matrix of responses by Observer 6 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		8																				
	B		6	2																			
	C		5	1		1		1															
f	A			8																			
	B		1	7																			
	C			8																			
z	A	1			7																		
	B				8																		
	C				8																		
ð	A		2			4	2																
	B		1			4	3																
	C		1			7																	
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 22. Matrix of responses by Observer 7 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A	1	5					1	1														
	B		5	3																			
	C		4	2				2															
f	A		2	3				3															
	B			8																			
	C		2	6																			
z	A	1			6										1								
	B				8																		
	C				7		1																
ð	A		2			3	3																
	B		1			1	5						1										
	C					4	3								1								
v	A					1	7																
	B					1	7																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								7	1													
	B								8														
	C							3	5														
k	A								1	7													
	B								1	7													
	C									8													
b	A										8												
	B										7	1											
	C										8												
d	A										1	7											
	B											8											
	C										1	7											
g	A											1	7										
	B												8										
	C						1						7										

Figure 23. Matrix of responses by Observer 7 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		8																				
	B		7	1																			
	C		7	1																			
f	A		2	6																			
	B		1	7																			
	C		1	7																			
z	A				8																		
	B				8																		
	C				8																		
ð	A					6	2																
	B					7	1																
	C					8																	
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 24. Matrix of responses by Observer 8 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	7	1																				
θ	A		6	1					1														
	B		4	3						1													
	C			2	4			2															
f	A		1	5				1		1													
	B			7																		1	
	C		1	7																			
z	A				6	2																	
	B				8																		
	C				2	3	1					1					1						
ð	A					6	2																
	B					5	2												1				
	C					5									1	2							
v	A						8																
	B					1	7																
	C					1	6									1							
p	A							7		1													
	B							6		2													
	C							8															
t	A								7	1													
	B								8														
	C							2	6														
k	A								2	6													
	B								1	7													
	C									8													
b	A						1				7												
	B						2				6												
	C					1	3				4												
d	A											8											
	B					1						7											
	C					1	1					4	2										
g	A								1				7										
	B												7						1				
	C												8										

Figure 25. Matrix of responses by Observer 8 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli		s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		8																				
	B		6	2																			
	C		8																				
f	A		1	7																			
	B		2	6																			
	C			8																			
z	A				8																		
	B				8																		
	C				8																		
ð	A					3	5																
	B					4	4																
	C					7	1																
v	A						8																
	B						8																
	C					1	7																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										8												
	C										8												
d	A											8											
	B										1	7											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 26. Matrix of responses by Observer 9 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	7		1																			
	B	8																					
	C	7	1																				
θ	A		4	2						2													
	B		3	3					1	1													
	C		4	2					1	1													
f	A			6						2													
	B			8																			
	C		1	7																			
z	A				7	1																	
	B				8																		
	C				8																		
ð	A					3	5																
	B					3	4					1											
	C					4	4																
v	A						7				1												
	B					2	6																
	C				1	2	5																
p	A							8															
	B					1		7															
	C							8															
t	A								7	1													
	B								8														
	C								8														
k	A								1	7													
	B							1	1	6													
	C								2	6													
b	A						2				6												
	B						1				5		2										
	C						2				6												
d	A											8											
	B						1					6	1										
	C										2	6											
g	A											2	6										
	B												8										
	C						1					1	6										

Figure 27. Matrix of responses by Observer 9 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	8																					
θ	A		7	1																			
	B		6	2																			
	C		7	1																			
f	A		1	7																			
	B			8																			
	C			8																			
z	A				8																		
	B				8																		
	C				8																		
ð	A		1			4	3																
	B					4	4																
	C					8																	
v	A						8																
	B						8																
	C						8																
p	A							8															
	B							8															
	C							8															
t	A								8														
	B								8														
	C								8														
k	A									8													
	B									8													
	C									8													
b	A										8												
	B										7			1									
	C										8												
d	A											8											
	B											8											
	C											8											
g	A												8										
	B												8										
	C												8										

Figure 28. Matrix of responses by Observer 10 to undistorted consonants.

Letters A, B, and C identify talkers.

Number of Responses

Consonants Used as Stimuli

Stimulus	Talker	s	θ	f	z	ð	v	p	t	k	b	d	g	omit	l	r	w	m	n	ŋ	ʃ	h	tr
s	A	8																					
	B	8																					
	C	3	5																				
θ	A		6					1	1														
	B		2					1	2	3													
	C		4					3			1												
f	A		1	3				4															
	B			5				3															
	C		2	3				2		1													
z	A				3	5																	
	B				8																		
	C				2	4	2																
ð	A		2				6																
	B		3			1	4																
	C					6	1									1							
v	A						8																
	B					2	6																
	C					1	6									1							
p	A							8															
	B							8															
	C							8															
t	A								7	1													
	B								8														
	C							2	6														
k	A							1	1	6													
	B							1	2	5													
	C								1	7													
b	A							1			7												
	B										7		1										
	C										8												
d	A											8											
	B											7	1										
	C										1	5	2										
g	A									1			7										
	B												8										
	C						1						7										

Figure 29. Matrix of responses by Observer 10 to consonants distorted by infinite peak clipping.

Letters A, B, and C identify talkers.

BIBLIOGRAPHY

Ahmend, R., and Fatehchand, R., Effect of sample duration on the articulation of sounds in normal and clipped speech, J. Acous Soc. Am., 31, 1959, 1022-1029.

Black, J. W., Accompaniments of word intelligibility, J. Speech Hearing Dis., 17, 1952, 409-418.

Black, J. W., and Hixson, W. C., Number of axis crossings and intelligibility of speech, J. Acous. Soc. Am., 31, 1959, 1384-1385.

Brandt, A. E., Forms of analysis for either measurement or enumeration data amenable to machine methods, Proceedings, Computation Seminar, December 1949, N. Y., International Business Machines Corporation.

Davis, H., Stevens, S. S., Nichols, R. H., Jr., Hudgins, C. V., Marquis, R. J., Peterson, G. E., and Ross, D. A., Hearing Aids: An Experimental Study of Design Objectives. Cambridge, Massachusetts: Harvard University Press, 1947.

Dukes, J. M. C., The effect of severe amplitude limitation on certain types of random signal: a clue to the intelligibility of 'infinitely' clipped speech, Monograph No. 111R, Proc. Instn. of Elect. Engrs., 1954, 88-97.

Egan, J. P., Articulation testing methods, The Larnygoscope, 58, 1948, 955-991.

Harris, J. D., Combinations of distortion in speech, A.M.A. Archives of Otolaryngology, 72, 1960, 227-232.

Harris, K. S., Cues for the discrimination of American English fricatives in spoken syllables, Language and Speech, 1, 1958, 1-7.

Hirsh, I. J., Reynolds, E. G., and Joseph, M., Intelligibility of different speech materials, J. Acous. Soc. Am., 26, 1954, 530-538.

Jacobson, R., Fant, C. G. M., and Halle, M., Preliminaries to speech analysis, Tech. Rept. No. 13, Acoustics Lab., M.I.T., 1952.

Kryter, K. D., Licklider, J. C. R., and Stevens, S. S., Premodulation clipping in AM voice communication, J. Acous. Soc. Am., 19, 1947, 125-131.

Licklider, J. C. R., The intelligibility of amplitude-dichotomized, time-quantized speech waves, J. Acous. Soc. Am., 22, 6, 1950, 820-823.

Licklider, J. C. R., Bindra, D., and Pollack, I., The intelligibility of rectangular speech-waves, Am. J. Psychol., 61, 1948, 1-20.

Licklider, J. C. R., and Pollack, I., Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech, J. Acous. Soc. Am., 20, 1948, 42-51.

Littwin, S., Ed., Pulse Generators in Industrial Electronics. New York: John F. Rider Publisher, 1964.

Miller, G. A., Language and Communication, New York: McGraw-Hill, 1951.

Miller, G. A., Heise, G., and Lichten, W., The intelligibility of speech as a function of the context of the test materials, J. Exp. Psychol., 41, 1951, 329-335.

Miller, G. A., and Nicely, P. E., Analysis of perceptual confusions among some English consonants, J. Acous. Soc. Am., 27, 1955, 338-352.

Pollack, I., and Pickett, J. M., Intelligibility of peak-clipped speech at high noise levels, J. Acous. Soc. Am., 31, 1959, 14-16.

Snedecor, G. W., Statistical Methods, Fifth Edition, Ames, Iowa: The Iowa State University Press, 1956.

Southworth, G. R., Speech digitizing techniques, J. Audio Engr. Soc., 11, 4, October 1963, 349-352.

Velichkin, A. I., Amplitude clipping of speech, Soviet Physics-Acoustics, 8, 2, Oct.-Dec. 1962, 130-134.

Vilbig, F., and Haase, K. H., Theoretical investigations to reduce harmonic distortions in a clipping process, J. Acous. Soc. Am., 29, 1957, 776.

BIOGRAPHICAL DATA

Maurice Joseph was born November 25, 1926, at Astoria, New York. In February 1944 he was graduated from William Cullen Bryant High School in Astoria. From 1944 to 1946 he served in the United States Army. Following his discharge he returned to school and subsequently received the degree of Bachelor of Arts from Queens College, Flushing, New York, in February 1950. He was a graduate assistant at the University of Pittsburgh and received the degree of Master of Science in August 1952. After two years as a research assistant at Central Institute for the Deaf in St. Louis, Missouri, Mr. Joseph taught for ten years at Tulane University Medical School. In 1964 he was enrolled at the University of Florida where he has continued to work toward the degree of Doctor of Philosophy.

Mr. Joseph has published articles, as junior author, in the _Journal of the Acoustical Society of America_ and in the _Southern Medical Journal_. He was granted a Special Rehabilitation Research Fellowship by the Vocational Rehabilitation Administration.

Maurice Joseph is married to the former Pauline Julia Bruckenstein and is the father of two children, Laura and David. He is a member of the American Speech and Hearing Association, the American Association for the Advancement of Science, the American Association of University Professors, and Sigma Alpha Eta.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of that committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council, and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 17, 1966

E. Ruffin Jones

Dean, College of Arts and Sciences

Dean, Graduate School

SUPERVISORY COMMITTEE:

Donald Dew

Chairman

Harry Hollien

George T. Singleton

E. P. Horne

John F. Brandt

Paul Moore

Richard J. Anderson